

Progress of the WDS/RDA Working Group on Assessment of Data Fitness

Robert R. Downs¹,

¹ rdowns@ciesin.columbia.edu

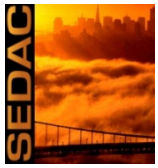
NASA Socioeconomic Data and Applications Center (SEDAC)
Center for International Earth Science Information Network (CIESIN)
The Earth Institute, Columbia University

ESIP 2018 Winter Meeting
Bethesda, Maryland

Session: Information Quality – Progress on Many Fronts
Thursday, 11 January 2018 11:00 a.m. - 12:30 p.m.



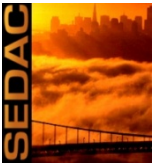
WDS/RDA WG on Assessment of Data Fitness



- Scope: Data quality identification, assessment, and understanding
- Co-Chairs: Helena Cousijn, Michael Diepenbroek, Claire Austin, Jonathan Petters
- Sponsoring Igs:
 - WDS/RDA Publishing Data Interest Group
 - WDS/RDA Certification of Digital Repositories Interest Group
- <https://www.rd-alliance.org/groups/assessment-data-fitness-use>
- RDA WGs engage in specified activities for 18 months



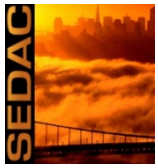
Progress on Selected WG Activities



- International Research Data Management (IRiDiuM) Glossary
 - Identification of terms and definitions
- Compiling Criteria for Data Fitness for Use
 - Gathering relevant data fitness criteria



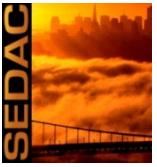
IRiDiuM Glossary



- International Research Data Management (IRiDiuM) Glossary
 - Practical Reference for data management
 - Definitions for standards terminology
- IRiDiuM Glossary WG
 - CASRAI (Consortia Advancing Standards in Research Administration Information)
 - CODATA (International Council for Science Committee on Data for Science and Technology)
 - Research Data Canada
- For use across communities
 - Research management, policy, authorship, support, publication, sponsorship
- Living document
 - Terms may be refined or evolve over time
 - http://dictionary.casrai.org/Category:Research_Data_Domain



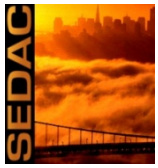
IRiDiuM Glossary WG Progress



- Phase 1
 - Identification of terms
 - Definition of terms
 - Open Review (completed Spring 2017)
- Phase 2
 - Assess comments from previous open review
 - Finalize completed terms for inclusion in online glossary (where?)
 - Propose new terms to be defined
 - Consider editorial improvements to existing terms



Compiling Criteria for Data Fitness for Use



- WDS/RDA Assessment of Data Fitness for Use
 - Criterion of Fitness for Use (being developed)
- FAIR Principles
 - Findable, Accessible, Interoperable, Reproducible
- GEOSS DMP Labels
 - Group on Earth Observations System of Systems Data Management Principles
- ESIP Information Quality Cluster aspects of data quality
 - Scientific Quality, Product Quality
- Trusted Data Repositories Proxy
 - CoreTrustSeal (<https://www.coretrustseal.org/>)

GEOSS Data Management Principles



Discoverability

DMP-1: Metadata for Discovery



Accessibility

DMP-2: Online Access



Usability

DMP-3: Data Encoding



DMP-4: Data Documentation



DMP-5: Data Traceability



DMP-6: Data Quality-Control



Preservation

DMP-7: Data Preservation



DMP-8: Data and Metadata Verification



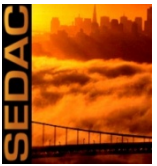
Curation

DMP-9: Data Review and Reprocessing



DMP-10: Persistent and Resolvable Identifiers





<https://www.rd-alliance.org/groups/assessment-data-fitness-use>